

DOCUMENT RESUME

ED 138 240

IR 004 620

AUTHOR Atherton, Pauline
TITLE Subject Access Project. Third Quarterly Report.
INSTITUTION Syracuse Univ., N.Y. School of Information
Studies.
PUB DATE Mar 77
NOTE 6p.; For related documents, see IR 004 775 and ED 129
229

EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
DESCRIPTORS *Cataloging; *Data Bases; *Experimental Programs;
*Indexing; Information Retrieval; Library Research;
On Line Systems; Subject Index Terms; University
Libraries
IDENTIFIERS Machine Readable Cataloging; MARC; *Subject Access

ABSTRACT

This third quarterly report for the period January to March 1977 describes the production schedule, records, and estimated costs and times in creating the Subject Access Project data base. Plans for on-line use of the data base and search strategy design are outlined. A table of specifications for preparing the data base for on-line searching is included, and appended is an augmented subject description for the record. (KP)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

ED138240

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

SYRACUSE UNIVERSITY

SCHOOL OF INFORMATION STUDIES

SUBJECT ACCESS PROJECT

THIRD QUARTERLY PROGRESS REPORT

March, 1977

PROJECT STAFF

Pauline Atherton - Project Director
B.K.L. Genova - Asst. Project Director
Bette Brindle - Research Associate
Barbara Settel - Research Assistant
Cheryl Willson - Secretary

Summary of Work in Progress

- (1) Creating the Data Base
- (2) MARC-SAP Data Base - Check in OCLC
- (3) BRS, LRS, and SDC On-Line Searches
- (4) Plan for On-Line Use of MARC-SAP Data Base

IR004 620

I. Creating the Data Base

This has been a time of production schedules for getting our data base ready for on-line searching.

The tasks involved five distinct operations:

- (1) Selecting entries from tables of contents and/or indexes, followed by a review of the selections made and coding the entries for input.
- (2) Keying the selections for each record, thereby creating a file of machine-readable data. On-line input via the SOS text-editing computer program on the PDP-10 at Syracuse University was used for this purpose.
- (3) Proofreading the listings of these keyed-in records which we obtained off the high-speed printer.
- (4) Correcting any errors by on-line text editing.
- (5) Merging our file of selections with the cataloging information file we received from the University of Toronto.

Our production records to date are as follows (as of March 1, 1977):

<u>Task</u>	<u>Done</u> (N = 2,007)
1) Selecting	1,772
2) Keying	939
3) Proofreading	425
4) Correcting	---
5) Merging	---

Our current estimates of costs and times (per unit) for each operation are as follows:

<u>Task</u>	<u>Average Staff Time/Unit</u> (no supervisory time incl.)	<u>Costs/Unit*</u>
1) Selecting	15 min.	\$.80
2) Keying	20-25 min.	1.50
3) Proofreading	5 min.	.25
4) Correcting	5 min.	.40
5) Merging	?	?

* (Staff plus computer time, supplies, but no overhead or fixed costs amortized)

We plan to ship the first portion of our data base to System Development Corporation no later than March 31. We have been promised access to the data base on-line no later than one month after receipt of our tape, barring unforeseen difficulties.

Attached is an example of the records we are creating to augment the MARC-like catalog records we received from the University of Toronto.

II. MARC-SAP data Base -- Check in OCLC

We are checking to see how many of our titles are also in the OCLC files.

Our first search was in the field of Urban Planning (HT). Of 336 searched, 254 were found in OCLC (76%).

We think this makes our data base even more interesting to academic libraries who might be interested in searching our data base when it is on-line.

III. BRS, LRS, and SDC On-Line Searches

To familiarize staff with on-line searching, several trial searches have been run on ERIC and Psychological Abstracts data bases. Our plans for searching SDC/MARC-SAP on-line follow.

IV. Plans for On-Line Use of MARC-SAP Data Base

We have received a number of "real queries" from the University of Toronto, solicited by the Robarts Library Reference staff. Examination of these queries and our own interview of graduate students and faculty at several upstate New York campuses have helped us to design the searching strategies we will permit with our data base. Our specification to SDC of searchable and printable fields, and the flexibility built into the manipulation of our augmented subject descriptions are critical decisions.

Here is a table of our present specs for preparing the data base for on-line searching. Attached in the appendix is the catalog record and our augmented subject description for one record (case no.1473).

Preliminary SPECS for MARC-SAP/SDC On-Line Data Base

<u>MARC Tag</u>	<u>Printable</u>	<u>Searchable</u>	<u>Free Term</u> <u>(FT)</u>	<u>Index Term</u> <u>(IT)</u>
RSN	X	X	X	
CCS	X	X	X	
MEP	X	X	X	
TIL	X	X	X	
IMP	X			
COL	X			
SER	X	?		
NOB	X	?		
SAP	X ^b	X	X	X ^c
SUT	X	X	X ^a	X ^a
LCN	X	X	X	
SBN	X	X	X	
DDC	X	X	X	
Fixed Field #3:Date		X	X	
SAP Case No.	X	X	X	

-
- These subject headings will be bound (IT) as in card catalog as well as unbound (FT) as in a KWIC index.
 - This field will be printed in part as well as in full. Weighting will be based on number of pages available on the topic.
Terms with ten or more pages will be starred (*).
Terms with two or more locator sequences will be starred (*).
 - Terms will be bound as IT terms if they meet certain rules dictated by the punctuation marking their relationship. See separate page for examples and rules.

Augmented Subject Description for MARC

SAP UTLAS
case no. record no.
147300887937:

index: @ aerial perspective (p. 49-54) @ anxiety (p. 30-35) @ auditory cues
loudness and localization (p. 88-93) @ blindness and vertical orientation
(p. 101-105) @ communication (p. 78-100) @ constancy size (p. 54-66) @ danger
(p. 146-166) @ divers and diving ; hearing (p. 92-97) ; orientation (p. 101-105)
@ experience and navigation (p. 108-112) @ eye movements (p. 126-130) @ gases
(p. 26-30) @ hearing under water (p. 92-97) @ kinaesthetic cues (p. 101-106)
@ motivation (p. 156-161) @ movement ; and illusions (p. 125-145) ; of bodies
(p. 134-143) @ navigation (p. 108-121) ; learning (p. 108-112) @ orientation
(p. 101-124) @ performance limits to (p. 26-48) @ risks (p. 146-160) @ roads
and paths (p. 70-74) @ size errors and illusions (p. 54-65) @ slope perception
(p. 67-74) @ sound (p. 78-100) ; under water (p. 92-97) @ spatial perception
(p. 49-77 89-93 101-121) @ sports dangerous (p. 156-161) @ tactile kinaesthetic
cues (p. 101-105) @ vertical orientation (p. 101-108) @ vestibular cues and
system (p. 101-105) @ vision ; and spatial perception (p. 49-77) ; and movement
perception (p. 125-143) @ visual cues and vertical orientation (p. 101-105)
@ visual illusions ; size and distance (p. 49-66) ; slope (p. 67-74) ; movement
(p. 125-143) #

~

Underlined words are candidates for weighting because they meet the rule for inclusion: locator range is for ten or more pages or there are two or more locator ranges.

The form of the IT depends on the punctuation. Every IT must begin with a word or phrase which begins with @. Attached to it can be a word preceded by ; (semicolon). Words or phrases preceded by / (slash) are to be attached first to word or phrase in front of it which is preceded by ; and before it a word or phrase with @. Words or phrases preceded by % (percent) are to be attached first to word or phrase in front of it which is preceded by /, then word or phrase preceded by ; and then by ;. (There can be two or more phrases preceded by / or % or ;. In moving backward through the chain of punctuation, a duplicate punctuation symbol is to be ignored in building the chain of words for the IT phrase.